

RESULT 1

TMS5_HUMAN

ID TMS5_HUMAN STANDARD; PRT; 457 AA.
AC Q9H3S3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Transmembrane protease, serine 5 (EC 3.4.21.-) (Spinesin).
GN Name=TMPRSS5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=11741986; DOI=10.1074/jbc.M103645200;
RA Yamaguchi N., Okui A., Yamada T., Nakazato H., Mitsui S.;
RT "Spinesin/TMPRSS5, a novel transmembrane serine protease, cloned from
RT human spinal cord.";
RL J. Biol. Chem. 277:6806-6812(2002).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Brain-specific. Predominantly expressed in
CC neurons, in their axons, and at the synapses of motoneurons in the
CC spinal cord.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB028140; BAB20375.1; -.
DR HSSP; P00746; 1FDP.
DR MEROPS; S01.313; -.
DR Genew; HGNC:14908; TMPRSS5.
DR MIM; 606751; -.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPc; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS50287; SRCR_2; FALSE_NEG.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Serine protease; Signal-anchor;
KW Transmembrane.
FT DOMAIN 1 49 Cytoplasmic (Potential).
FT TRANSMEM 50 70 Signal-anchor for type II membrane

FT				protein (Potential).
FT	DOMAIN	71	457	Extracellular (Potential).
FT	DOMAIN	112	207	SRCR.
FT	DOMAIN	218	457	Serine protease.
FT	ACT_SITE	258	258	Charge relay system (By similarity).
FT	ACT_SITE	308	308	Charge relay system (By similarity).
FT	ACT_SITE	405	405	Charge relay system (By similarity).
FT	SITE	217	218	Cleavage (Potential).
FT	DISULFID	135	196	By similarity.
FT	DISULFID	148	206	By similarity.
FT	DISULFID	209	328	By similarity.
FT	DISULFID	243	259	By similarity.
FT	DISULFID	374	390	By similarity.
FT	DISULFID	401	429	By similarity.
FT	CARBOHYD	163	163	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	170	170	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	195	195	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	319	319	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	375	375	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	457 AA; 49574 MW; 64406AB4985A2651 CRC64;		

Query Match 100.0%; Score 2481; DB 1; Length 457;
 Best Local Similarity 100.0%; Pred. No. 1e-191;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSLMLDDQPPMEAQYAEEGPGPGIFRAEPGDQQHPISQAVCWRSMMRRC	60
Db	1	MSLMLDDQPPMEAQYAEEGPGPGIFRAEPGDQQHPISQAVCWRSMMRRC	60
Qy	61	AGVGSWLLVLYLCPAASQPISGTLQDEEITLSCSEASAEELLPALPKTVSFRINSEDFL	120
Db	61	AGVGSWLLVLYLCPAASQPISGTLQDEEITLSCSEASAEELLPALPKTVSFRINSEDFL	120
Qy	121	LEAQVRDQPRWLLVCHGWSPALGLQICWSLGHRLRTHHKGVNLTDIKLNSSQEFAQLSP	180
Db	121	LEAQVRDQPRWLLVCHGWSPALGLQICWSLGHRLRTHHKGVNLTDIKLNSSQEFAQLSP	180
Qy	181	RLGGFLEEAWQPRNNTSGQVVSRLCSECGARPLASRIVGGQSVAPGRWPWQASVALGFR	240
Db	181	RLGGFLEEAWQPRNNTSGQVVSRLCSECGARPLASRIVGGQSVAPGRWPWQASVALGFR	240
Qy	241	HTCGGSVLAPRWVVTAACMHSFRLARLSSWRVHAGLVSHSAVRPHQGALVERIIPHPLY	300
Db	241	HTCGGSVLAPRWVVTAACMHSFRLARLSSWRVHAGLVSHSAVRPHQGALVERIIPHPLY	300
Qy	301	SAQNHDYDVALLRLQTALNFSDTVGA VCLPAKEQHFPKGSRCVWVGWGHThPSHTYSSDM	360
Db	301	SAQNHDYDVALLRLQTALNFSDTVGA VCLPAKEQHFPKGSRCVWVGWGHThPSHTYSSDM	360
Qy	361	LQDTVVPFLFSTQLCNSSCVYSGALTTPRMLCAGYLDGRADACQGDSSGGPLVCPDGD TWRLV	420
Db	361	LQDTVVPFLFSTQLCNSSCVYSGALTTPRMLCAGYLDGRADACQGDSSGGPLVCPDGD TWRLV	420
Qy	421	GVVSWGRACAEPNHPGVYAKVAEFLDWIHDTAQDSLL	457
Db	421	GVVSWGRACAEPNHPGVYAKVAEFLDWIHDTAQDSLL	457

Database : UniProt_02:*
 1: uniprot_sprot:*
 2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2481	100.0	457	1	TMS5_HUMAN	Q9h3s3 homo sapien
2	1956	78.8	445	2	Q8CJ17	Q8cj17 rattus norv
3	1944	78.4	455	2	Q8CDR0	Q8cdr0 mus musculu
4	1941	78.2	455	1	TMS5_MOUSE	Q9er04 mus musculu
5	1642	66.2	371	2	Q8CJ16	Q8cj16 rattus norv
6	697.5	28.1	490	2	Q6P7D7	Q6p7d7 rattus norv
7	697.5	28.1	490	2	AAH61712	Aah61712 rattus no
8	696.5	28.1	453	1	TMS3_MOUSE	Q8k1t0 mus musculu
9	696	28.1	454	1	TMS3_HUMAN	P57727 homo sapien
10	696	28.1	490	2	Q7TN04	Q7tn04 mus musculu
11	695.5	28.0	453	2	Q812A6	Q812a6 mus musculu
12	691	27.9	490	1	TMS2_MOUSE	Q9jiq8 mus musculu
13	689.5	27.8	453	2	Q6ZMC3	Q6zmc3 homo sapien
14	689.5	27.8	453	2	AAQ88823	Aaq88823 homo sapi
15	689.5	27.8	453	2	BAD18806	Bad18806 homo sapi
16	677.5	27.3	490	2	Q920K3	Q920k3 rattus norv

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%	Query						
No.	Score	Match	Length	DB	ID		Description		
1	2481	100.0	457	9	US-09-888-615-110	<i>had date</i>	Sequence 110, App		
2	2481	100.0	457	10	US-09-981-151A-71	<i>had date</i>	Sequence 71, Appl		
3	2441.5	98.4	472	10	US-09-981-151A-27	<i>10/7/00</i>	Sequence 27, Appl		
4	2280	91.9	480	9	US-09-820-893-108		Sequence 108, App		
5	2280	91.9	480	15	US-10-607-565-108		Sequence 108, App		
6	2234	90.0	414	9	US-09-820-893-69		Sequence 69, Appl		

Database : Issued_Patents_AA:*
 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Match	Length	DB	ID	Description	
1	687.5	27.7	454	3	US-09-518-046-2	Sequence 2, Appli	
2	651.5	26.3	492	3	US-09-342-749-2	Sequence 2, Appli	
3	651.5	26.3	492	4	US-09-691-840-2	Sequence 2, Appli	
4	649.5	26.2	492	4	US-09-685-166A-895	Sequence 895, App	
5	649.5	26.2	492	4	US-09-879-792-14	Sequence 14, Appl	
6	649.5	26.2	492	4	US-09-679-426-895	Sequence 895, App	
7	644.5	26.0	562	4	US-09-879-792-12	Sequence 12, Appl	
8	633.5	25.5	477	4	US-10-177-661-2	Sequence 2, Appli	
9	624	25.2	446	4	US-10-177-661-4	Sequence 4, Appli	
10	619.5	25.0	416	2	US-09-000-846-2	Sequence 2, Appli	
11	617.5	24.9	418	4	US-10-177-661-6	Sequence 6, Appli	
12	604.5	24.4	417	4	US-09-820-002-4	Sequence 4, Appli	
13	592.5	23.9	283	3	US-08-807-151-1	Sequence 1, Appli	
14	592.5	23.9	283	3	US-09-478-957-1	Sequence 1, Appli	
15	588	23.7	376	4	US-09-820-002-2	Sequence 2, Appli	

US-09-820-893-108

Abandoned
all pair abandoned
CIP 7 w/o/oo/1722
510 108

Qy	37	SQAVCWRS	SMRRGCAVLGALGLLAGAGVGSWLLVLYLCPAASQPI	SGTLQDEEITLSCSEA	96			
Db	60	SQAVCWRS	SMRRGCAVLGALGLLAGAGVGSWLLVLYLCPAASQPI	SGTLQDEEITLSCSEA	119			
Qy	97	SAAEALL	PALPKTVSFRINSEDFLLEAQVRDQPRWLLV	CHEGWSPALGLQICWSLGHRL	156			
Db	120	SAAEALL	PALPKTVSFRINSEDFLLEAQVRDQPRWLLV	CHEGWSPALGLQICWSLGHRL	179			
Qy	157	THHKG	VNLTDIKLNSSQEFAQLSPRLGGFLEEAWQPRNNCT	SGQVVSLRCSECGARPLAS	216			
Db	180	THHKG	VNLTDIKLNSSQEFAQLSPRLGGFLEEAWQPRNNCT	SGQVVSLRCSECGARPLAS	239			
Qy	217	RIVGG	QSVAPGRWPWQASVALGFRHTCGGSVLAPRWVVTAAH	CMHSFRLARLSSWRVHAG	276			
Db	240	RIVGG	QSVAPGRWPWQASVALGFRHTCGGSVLAPRWVVTAAH	CMHSFRLARLSSWRVHAG	299			
Qy	277	LVSHS	AVRPHQGALVERIIPHPLYSAQNHDYDVALLRLQ	TALNFSDTVAVCLPAKEQHF	336			
Db	300	LVSHS	AVRPHQGALVERIIPHPLYSAQNHDYDVALLRLQ	TALNFSDTVAVCLPAKEQHF	359			
Qy	337	PKGSR	CWVSGWGH	THPSHTYSSDMLQDTVVP	LFSTQLCNSSCVYSGALT	TPRMLCAGYLDG	396	
Db	360	PKGSR	CWVSGWGH	THPSHTYSSDMLQDTVVP	LFSTQLCNSSCVYSGALT	TPRMLCAGYLDG	419	
Qy	397	RADAC	QGDSGGPLVCPD	GD	TWRLVGVVSWGRG	CAEPNHPGVYAKVAEFLDWIHDTA	QDSL	456
Db	420	RADAC	QGDSGGPLVCPD	GD	TWRLVGVVSWGRG	CAEPNHPGVYAKVAEFLDWIHDTA	QDSL	479
Qy	457	L	457					
Db	480	L	480					

Database : PIR_79:*
 1: pir1:*
 2: pir2:*
 3: pir3:*
 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	602.5	24.3	417	1	S00845	hepsin (EC 3.4.21.
2	599.5	24.2	416	1	S33777	hepsin (EC 3.4.21.
3	546.5	22.0	1019	1	A56318	enteropeptidase (E
4	524	21.1	638	1	KQHUP	plasma kallikrein
5	522.5	21.1	638	1	KQRTPL	plasma kallikrein
6	520.5	21.0	1034	1	A53663	enteropeptidase (E
7	519.5	20.9	638	1	KQMSPL	plasma kallikrein
8	515	20.8	1035	1	A43090	enteropeptidase (E
9	513.5	20.7	343	1	A57014	prostasin (EC 3.4.
10	510.5	20.6	625	1	KFHU1	coagulation factor
11	506	20.4	855	2	JC7731	membrane-bound arg
12	505	20.4	1113	2	JE0315	low-density lipopr
13	482.5	19.4	270	2	S56160	mast cell tryptase

Database : A_Geneseq_23Sep04:*
 1: geneseqp1980s:*
 2: geneseqp1990s:*
 3: geneseqp2000s:*
 4: geneseqp2001s:*
 5: geneseqp2002s:*
 6: geneseqp2003as:*
 7: geneseqp2003bs:*
 8: geneseqp2004s:*

Pred.. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	%		Query			Description
	Score	Match	Length	DB	ID	
1	2481	100.0	457	3	AAB11699	Aab11699 Human ser
2	2481	100.0	457	5	AAU82752	Aau82752 Amino aci
3	2441.5	98.4	472	7	ADJ38445	Adj38445 Human nov
4	2441.5	98.4	472	8	ADH41503	Adh41503 Novel hum
5	2439.5	98.3	472	5	ABG76906	Abg76906 Human hep
6	2415.5	97.4	568	7	ADI21256	Adi21256 Novel hum
7	2406	97.0	513	8	ADH41521	Adh41521 Novel hum
8	2400	96.7	513	8	ADH41501	Adh41501 Novel hum
9	2400	96.7	513	8	ADH41511	Adh41511 Novel hum
10	2400	96.7	513	8	ADH41515	Adh41515 Novel hum
11	2400	96.7	513	8	ADH41513	Adh41513 Novel hum
12	2400	96.7	513	8	ADH41505	Adh41505 Novel hum